

# Supernova burst trigger studies using machine learning for data selection

Guanqun Ge, Yeon-jae Jwa  
Columbia University



# Intro: ML based data selection

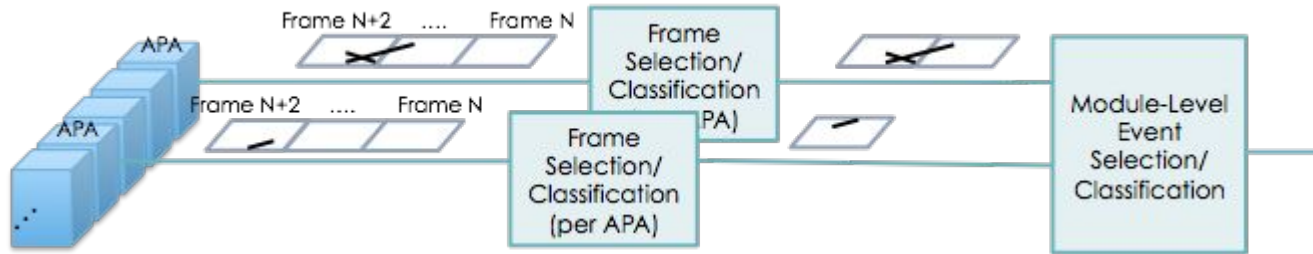
In [DUNE-doc-11311](#), we introduced a two-level triggering scheme for SN bursts using CNN image classification.

Two-level triggering scheme includes

- 1) Low-level: “APA-frame” CNN classification and selection,
- 2) Module-level: SN burst triggering using “APA-frame” selection coincidence (across multiple APA’s and during the expected SN burst duration).

# Intro: ML based data selection

Two-level triggering scheme:



**1. Low-level:**  
CNN-based  
APA-frame  
selection

**2. Module-level:**  
APA-frame coincidence  
across module and  
over 10 seconds  
(SN burst trigger)

# Intro: Low-level trigger

Applied on each **APA-frame** image (see backup slides for more details) independently.

**APA-frame = 960 collection wires x 2.25 ms (time-tick frequency: 2 MHz)**

Each APA-frame is classified by a CNN according to its contents, as

**SN/Low-energy interaction,**

**High-energy interaction** (inclusive of n-nbar, p-decay, atm. nu, cosmic), or

**None** (only radiological backgrounds and noise)

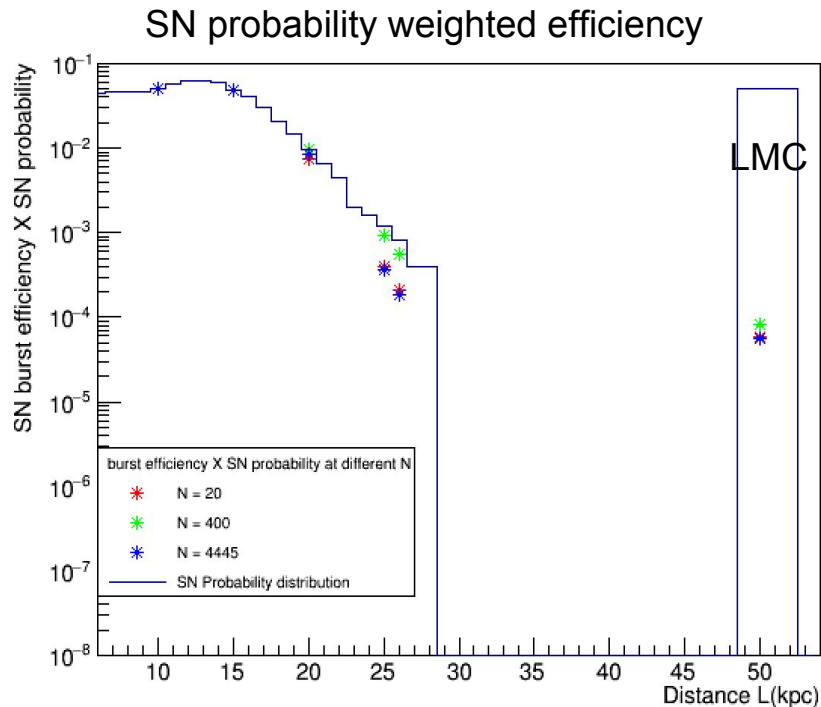
*Note:*

*This type of selection could in principle be applied at the front-end DAQ (given sufficient resources/power), or as a filtering stage after event building. For this talk, we consider the case of a front-end application of CNN classification and investigate subsequent burst trigger efficiency at module level. See docdb technote for more details: [DUNE-doc-11311](#)*

# Intro: Module-level trigger

- From CNN single frame selection, we find that  $1\text{E-}4$  background APA-frame reduction rate can be obtained, while SN interaction APA-frame efficiency is 69%. This means,
  - For any given SN burst, we could record each SN interaction APA-frame with ~69% efficiency, regardless of where the SN burst originated from. This can be done while keeping a steady-state data rate of 120 MB/s from the full 10kton module (dominated by the fake frame rate selection).
- We can introduce a **second level of trigger decision** involving an aggregated and prolonged scan of APA-frames searching for the signal of SN burst across the full 10kton module.

# ML based data selection: Performance of two-level trigger



Galactic coverage

$N$	Galactic Coverage	Uncertainty ( $10^{-5}$ )
20	0.981	5.96
200	0.997	1.25
400	0.998	0.835
500	0.998	0.895
600	0.998	0.974
4445	0.988	3.44

(Galactic coverage is defined as integral of the 'SN probability weighted efficiency' plot from 0 kpc to 28 kpc.)

# ML based data selection: Performance

With previous scheme ([DUNE-doc-11311](#)), we assume the CNN APA-frame selection efficiency is SN neutrino energy independent (we know it is not), and we also weigh each selected APA-frame equally in the burst trigger decision.

## Improvements with new approach (subsequent slides):

- Take into account energy vs. time dependency of the SN neutrinos, and corresponding **energy-dependent CNN efficiency for APA-frame selection**, for more realistic assumptions.
- Implement an “**Energy Boosted Decision**” algorithm in the burst triggering stage, where the (highest) deposited energy in the APA-frame is assumed to be available at the Module-level trigger stage. (Similar scheme suggested by Oxford and others in DAQ consortium for traditional approach.)

# The approach: Old & New

Low-level

CNN selects APA-frames  
with some efficiency  
(SN, background)

Module-level

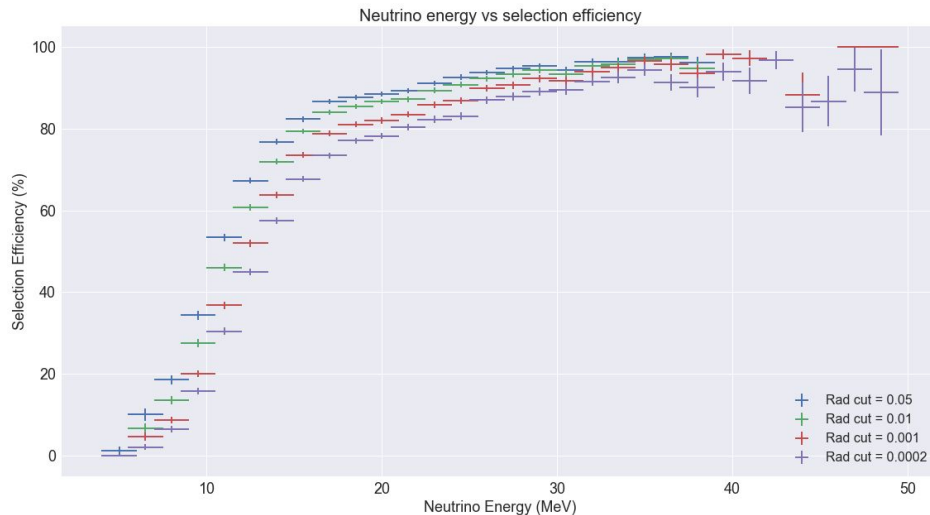
Burst tagging for 10s

Count the tagged  
APA-frames in N  
successive drift frames  
over 150 APAs  
→ multiplicity M

Find the maximum  
multiplicity  $M_{\max}$

Burst trigger on  $>M_{\text{cut}}$   
with low fake rate

Using '**flat**' efficiency from CNN so far  
does not take into account the energy  
dependency we have observed:



# The approach: Old & New

Low-level

CNN selects APA-frames  
with some efficiency  
(SN, background)

Using energy(true energy)-dependent  
efficiency of CNN for SN APA-frames

Module-  
level

Burst tagging for 10s

Count the tagged  
APA-frames in N  
successive drift frames  
over 150 APAs  
→ multiplicity M

Find the maximum  
multiplicity M<sub>max</sub>

Burst trigger on >M<sub>cut</sub>  
with low fake rate

“Boost” counting using a weight based on  
APA-frame energy from the output of  
subsequent algorithm (If algorithm quickly  
predicts the ‘energy’ of the single frame).

What is the ‘energy’ we are talking about?

- Should be an estimator of SN neutrino deposited energy.
- Assuming burst triggering will veto high E events.
- The energy of most energetic single MCTruth in single frame. (check needed if this corresponds to SN neutrino energy)

# Estimating energy at Low-/Module-level trigger stage:

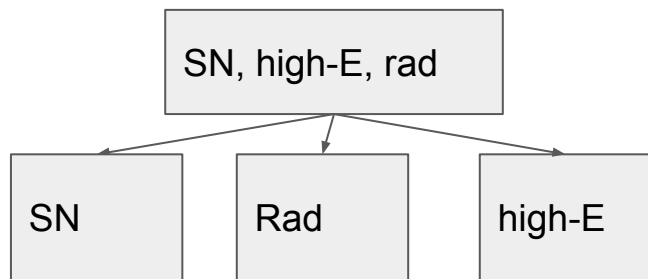
## Possible options:

1. Two sequential algorithms: APA-frame selector followed by energy classifier (CNN or otherwise), or,
2. Train the same CNN that selects APA-frames to also classify energy deposition in APA-frame

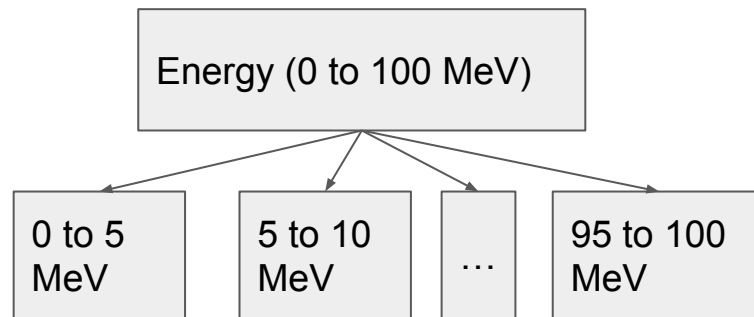
## Option 1:

CNN1: Physics process classifier: 3 scores with SN, high-E, radiological background only frames.

CNN2: Energy classifier: Predict the energy in the frame and provide as extra information handle in burst triggering

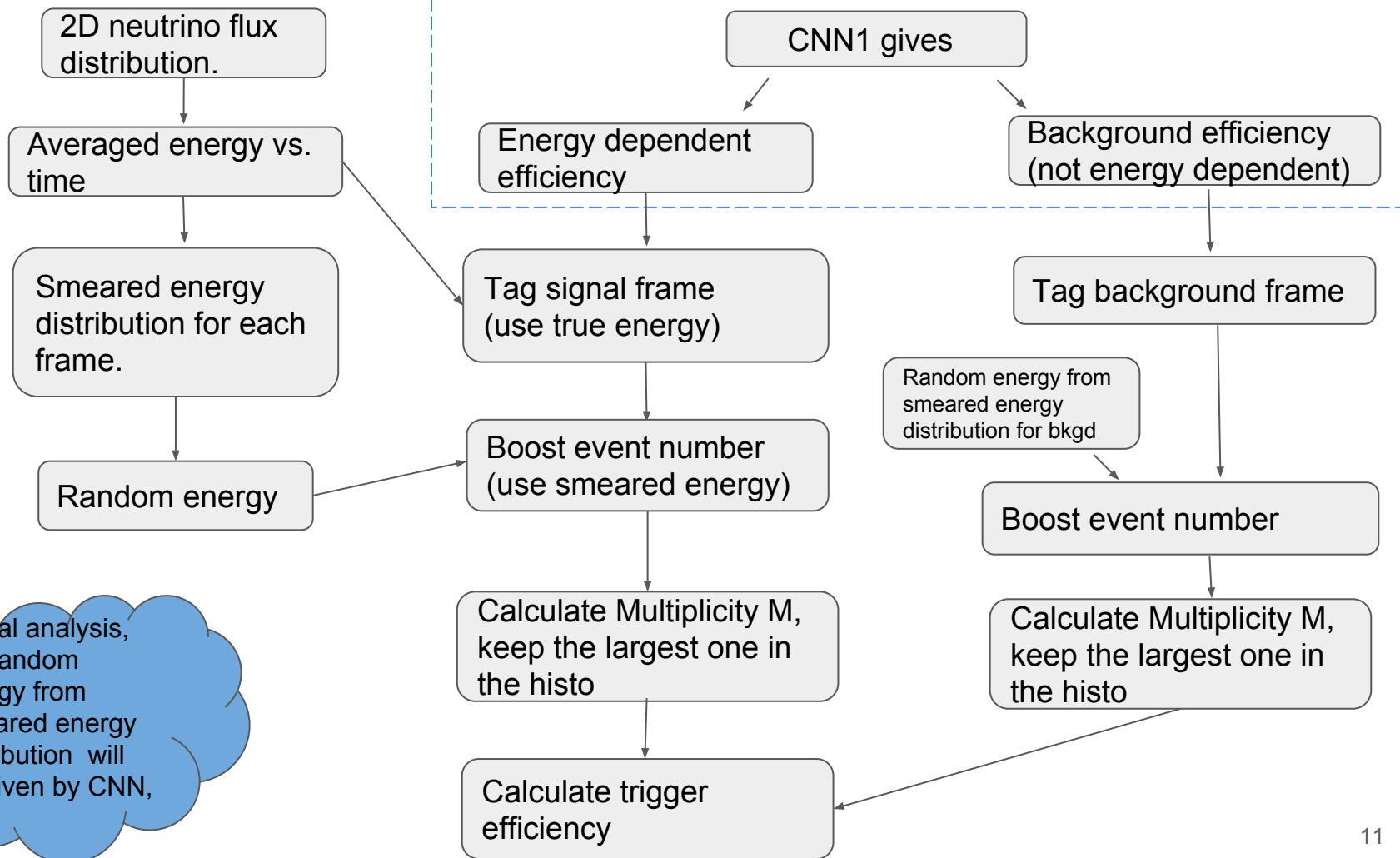


CNN1 classification goal



CNN2 classification goal  
(example binning)

## Details



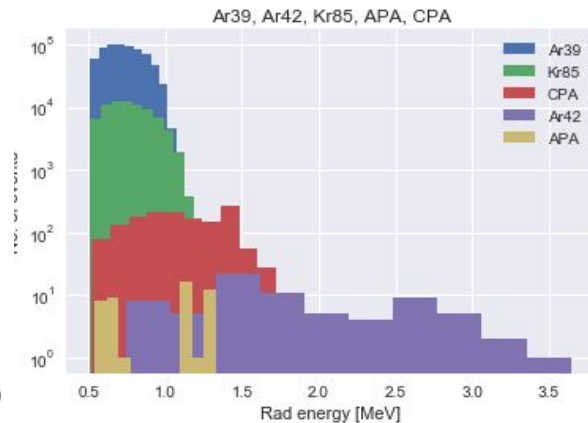
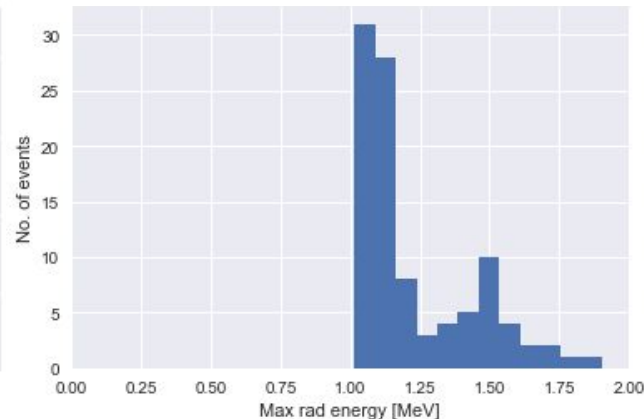
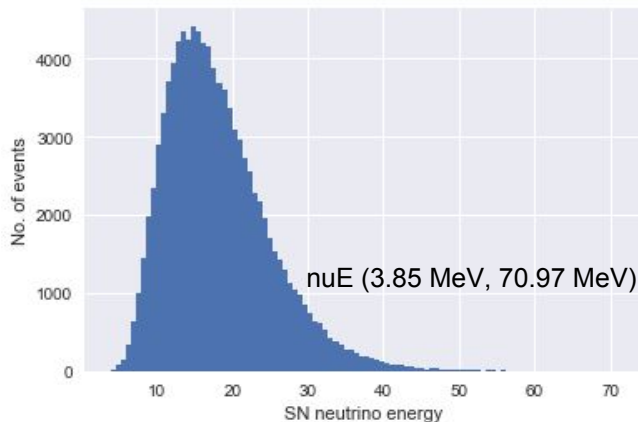
# Estimating energy with CNN2:

## Predicted deposited energy distributions:

Left: SN neutrino true energy;

Middle: the largest MCParticle energy in one Rad. frame;

Right: MCParticle energy distribution (Rad.)

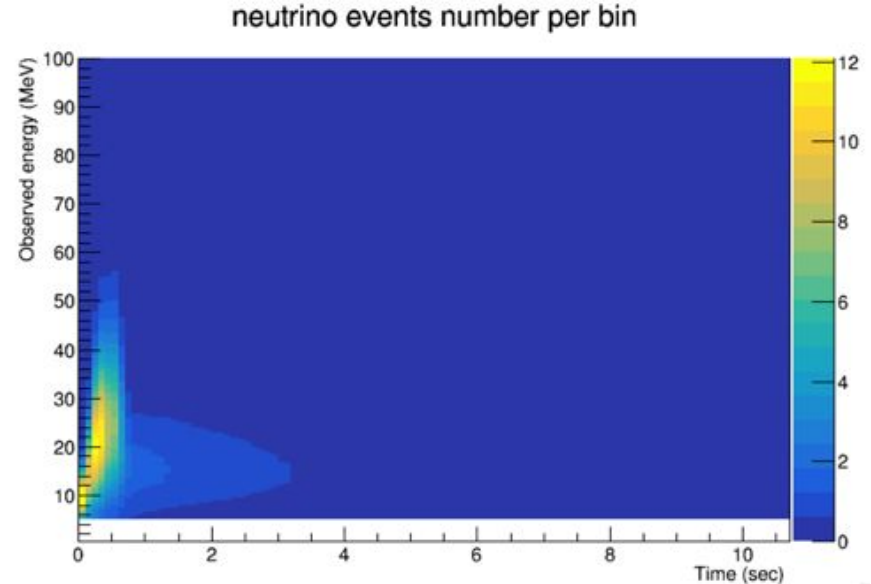
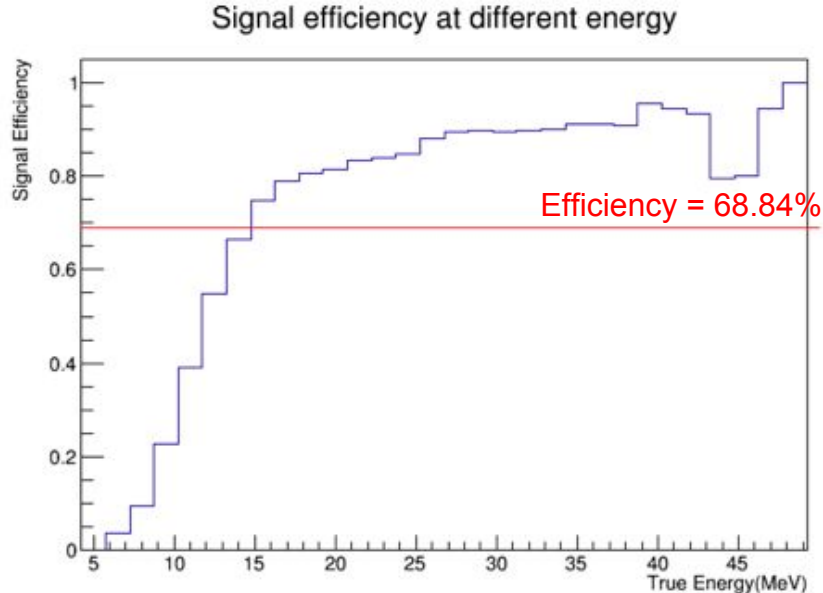


Radiological backgrounds and SN neutrinos have different deposited energy distributions.

If CNN2 can give quick and reliable energy classification, we can use the APA-frame weight in the multiplicity/burst trigger decision based on the classified energy.

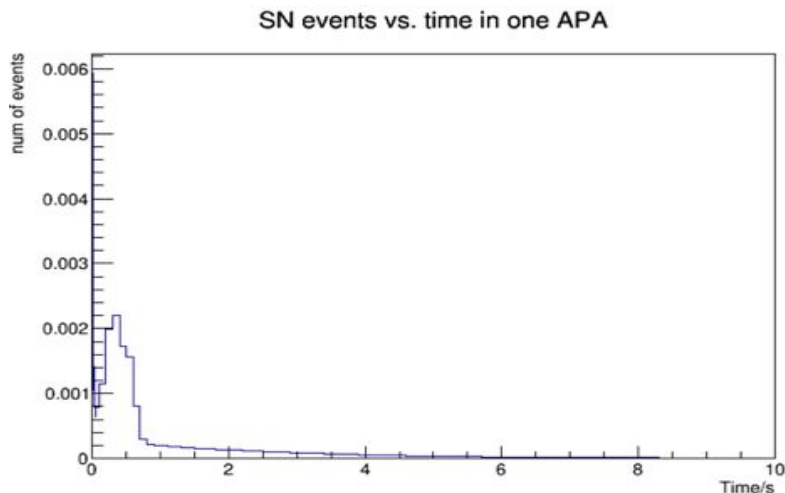
# 1. Energy-dependent CNN efficiency for APA-frame selection

- Energy dependent efficiency produced by CNN shows events with true energy larger than 15 MeV has efficiency larger than 69%.
- 2D SN timing profile (provided by Kate Scholberg) gives number of neutrino events with certain energy at different time.

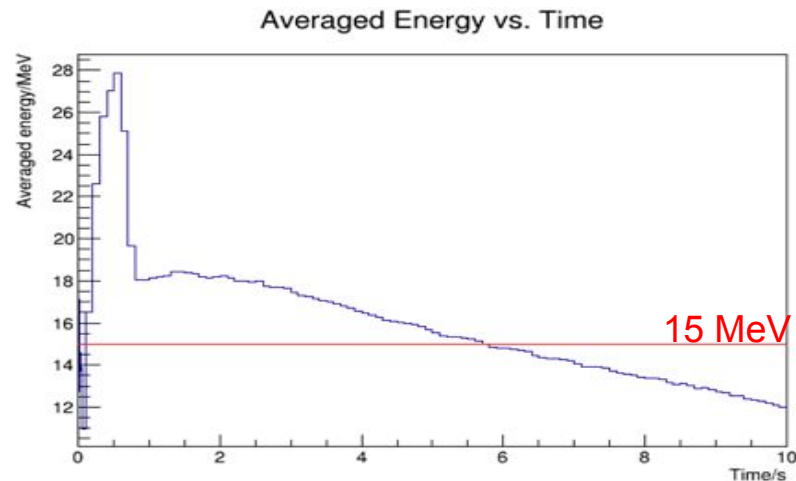


# 1. Energy-dependent CNN efficiency for APA-frame selection

- Integrating over energy gives 1D SN timing profile(number of events vs. time). Then it's applied to a certain distance to obtain event distribution per APA.
- Averaging over energy gives averaged energy distribution. For each drift frame (starting with  $t=0$ ), averaged true energy is found, then the corresponding CNN APA-frame selection efficiency(energy) is applied.



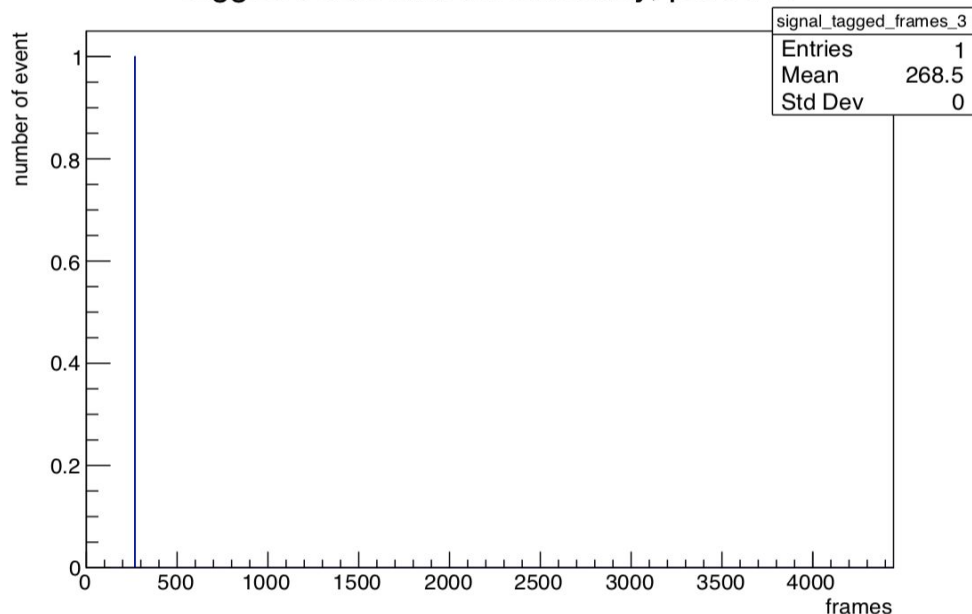
For supernova at distance  $L=10$  kpc



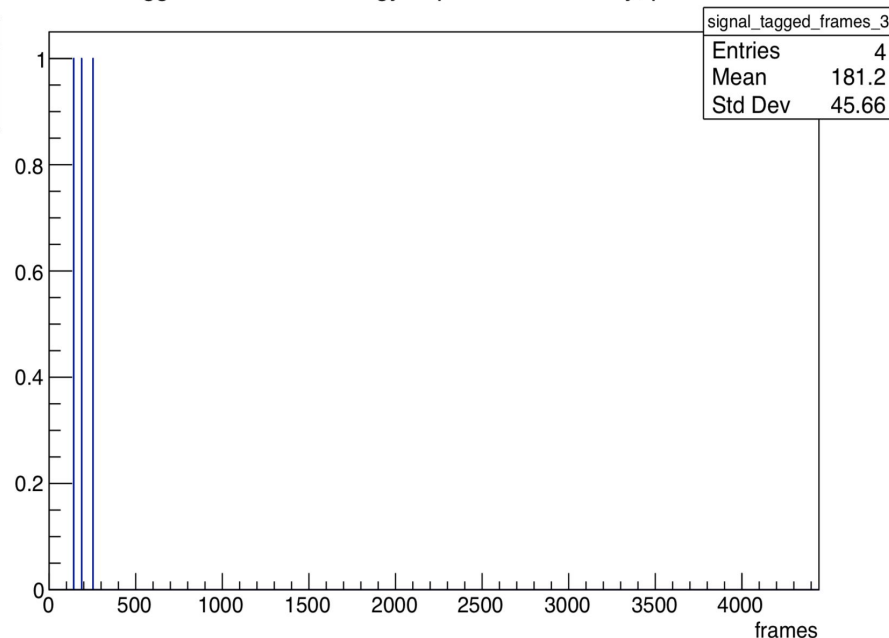
# 1. Energy-dependent CNN efficiency for APA-frame selection

- Comparison between result using energy-dependent efficiency and previous result

tagged frame with flat efficiency, per APA

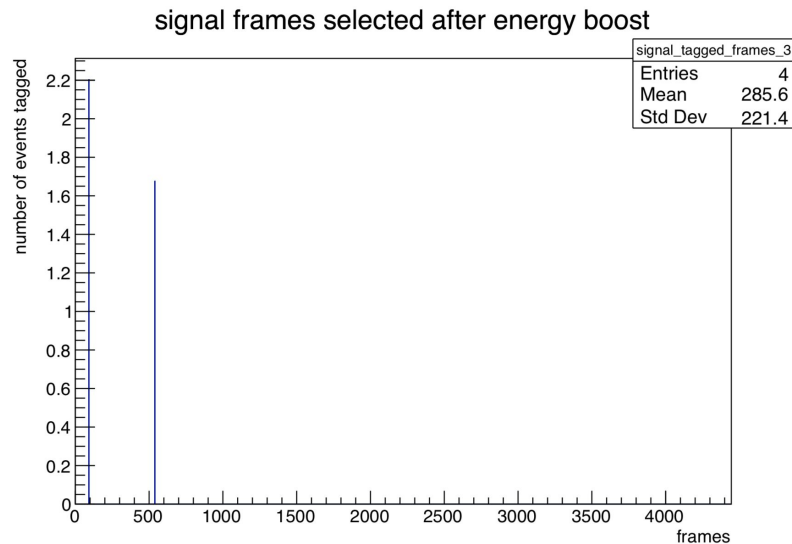
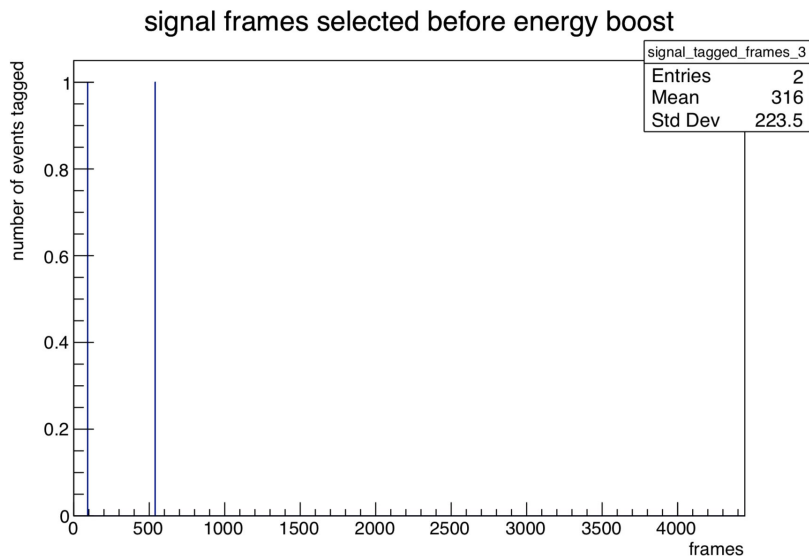


tagged frame with energy dependent efficiency, per APA



## 2. Energy boosted burst triggering

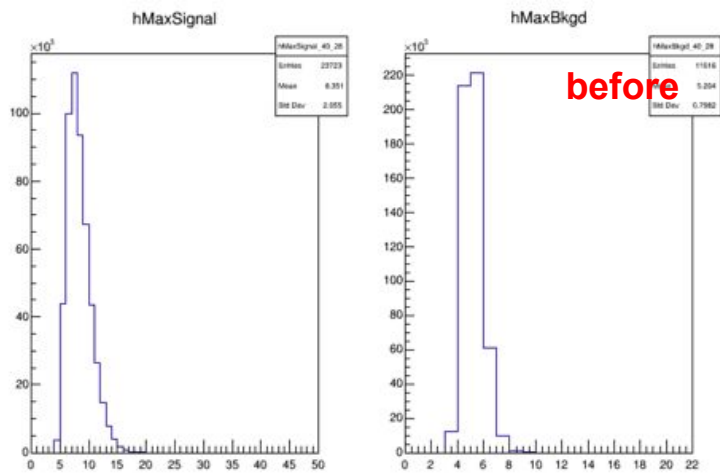
- For each tagged signal frame, vary the true averaged energy using a gaussian with mean=true energy, sigma=true energy\*10% to simulate resolution of detector/CNN.
- Draw energy from this (smeared) distribution; if the drawn energy is larger than 10 MeV, scale the event number in this frame linearly by a factor of (energy[MeV]/10).



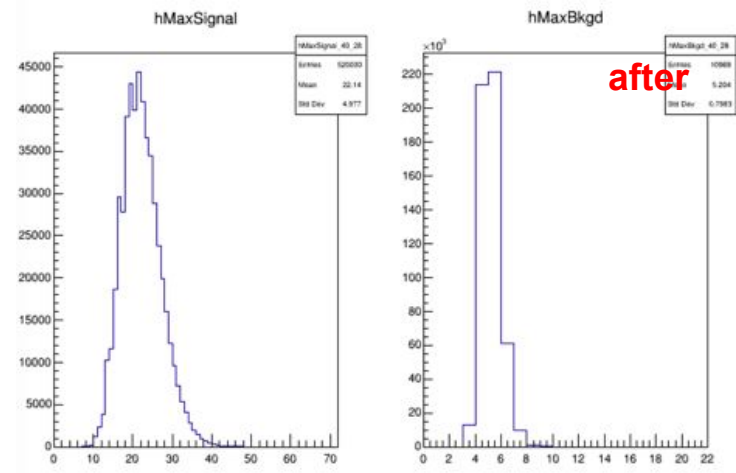
## 2. Energy boosted burst triggering

- For background, flat background efficiency is used to tag background frames. Energy boost in background frames will be applied once we have energy distribution for background events.
- After finishing energy boosting, multiplicity is calculated. Then a cut on multiplicity  $M_{\text{cut}}$  is applied to calculate burst trigger efficiency.

L=28 kpc, N=40

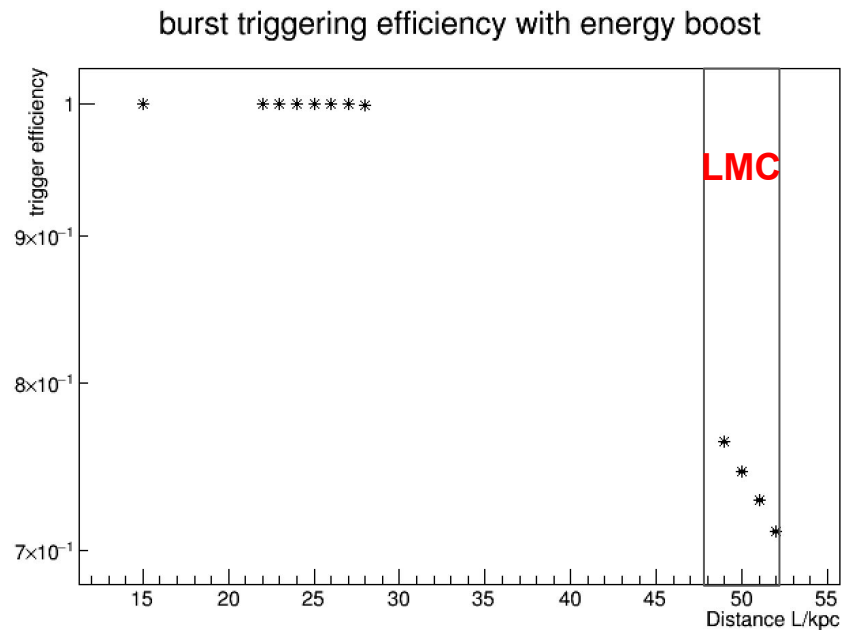
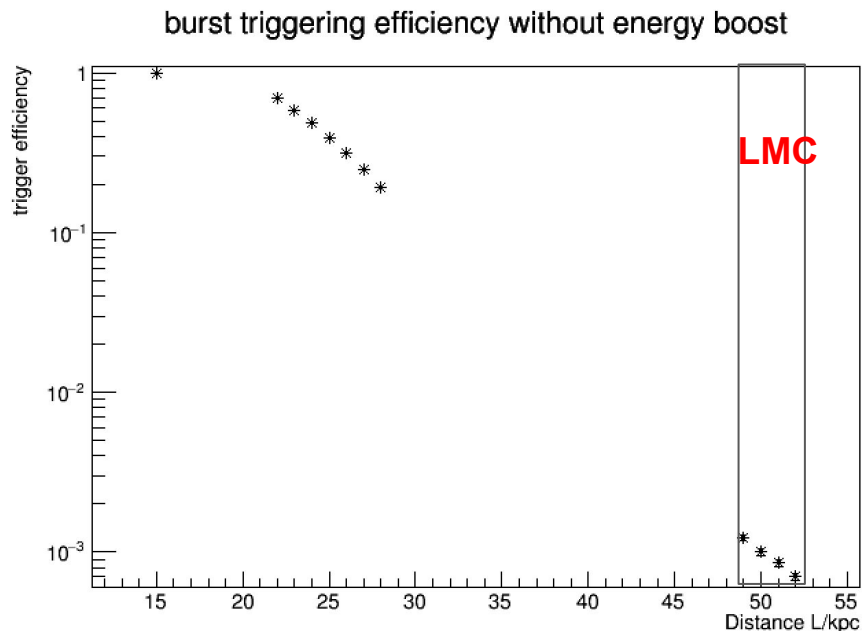


Triggering efficiency: 19.21%



Triggering efficiency: 99.93%

# Comparison between old & new scheme (for N=40): Burst trigger efficiency vs. distance:



# Comparison between old & new scheme (for N=40): Galactic coverage

	Old scheme	New scheme
Galactic coverage	0.988443	0.999999
LMC coverage	0.000946	0.737089

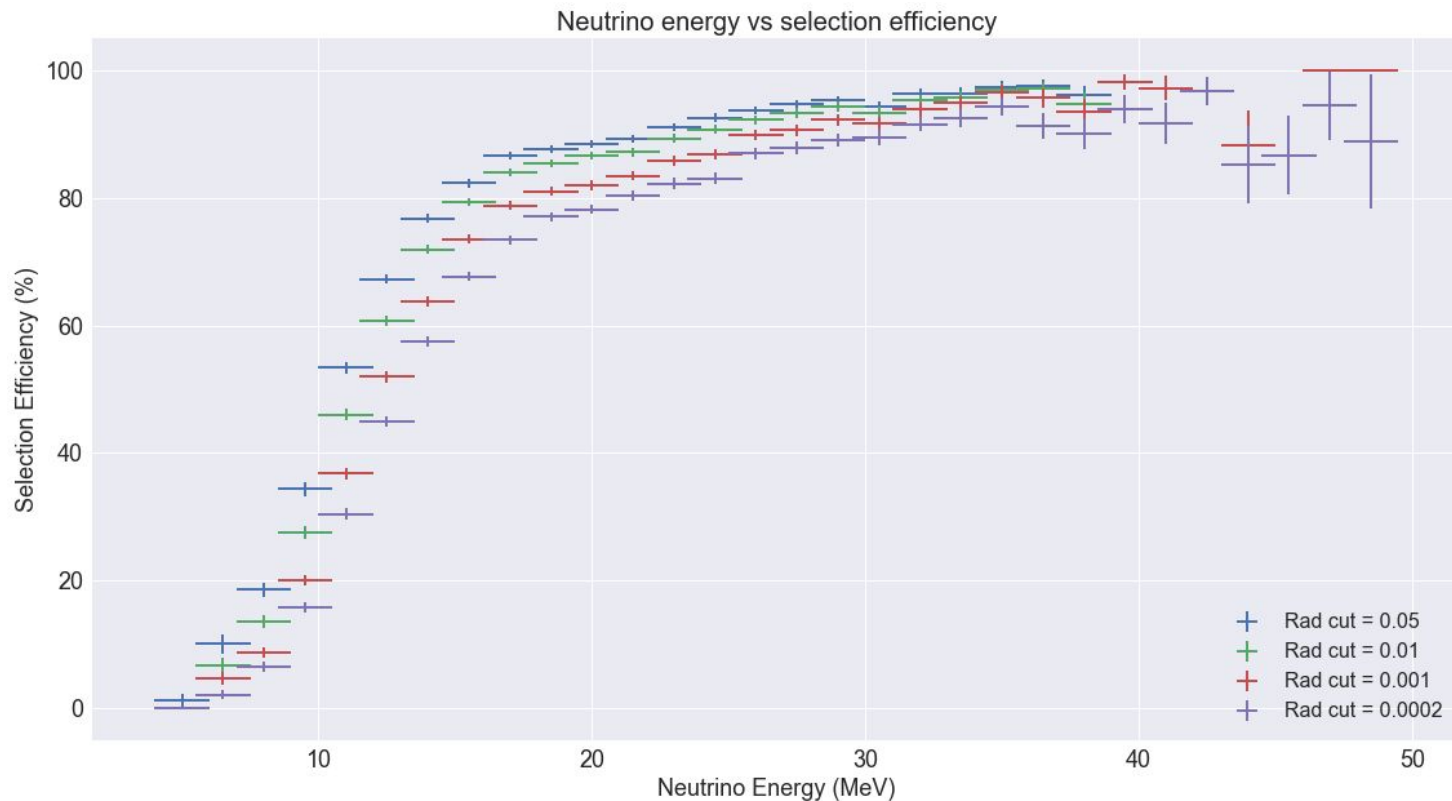
**Promising result with energy boosted burst trigger scheme:**  
**>99.9% galactic coverage and 73.7% LMC coverage**  
(assuming CNN2 can resolve energy to within 10% -- to be tested)

# Summary and Conclusions

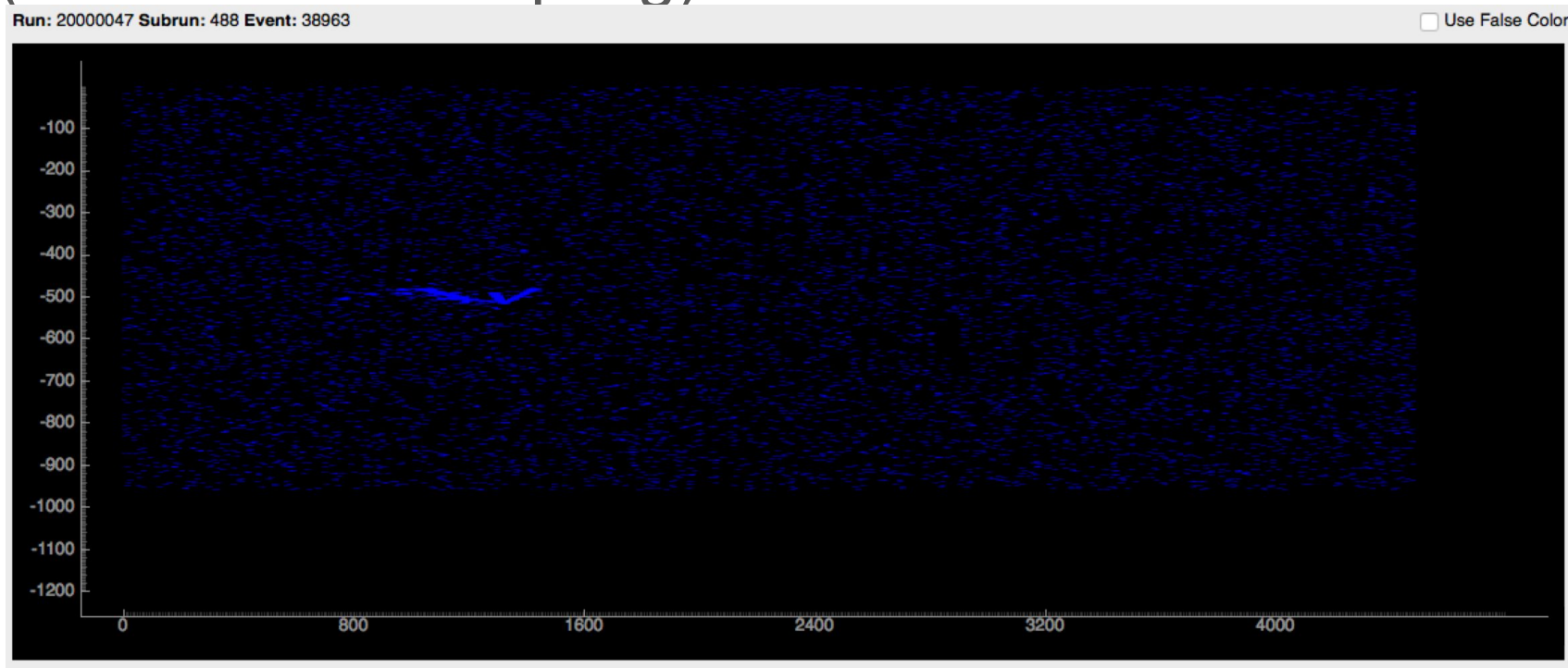
- Energy boost triggering yields higher selection efficiency for signals.
- **Future steps:**
  - CNN2 training (for determining selected APA-frame highest-deposited energy)
  - Apply energy boost weight to both SN and background (based on CNN2 results)
  - Try different energy boosting algorithms to maximize burst trigger efficiency
  - Study latency and power requirements for burst trigger decision
  - Consider a variety of SN time-energy evolution templates  
(how sensitive are we to time profile of SN neutrino flux?)

Back up

# Selection efficiency vs. Neutrino energy



E.g., high-energy (event) frame image input to CNN  
(before downsampling)

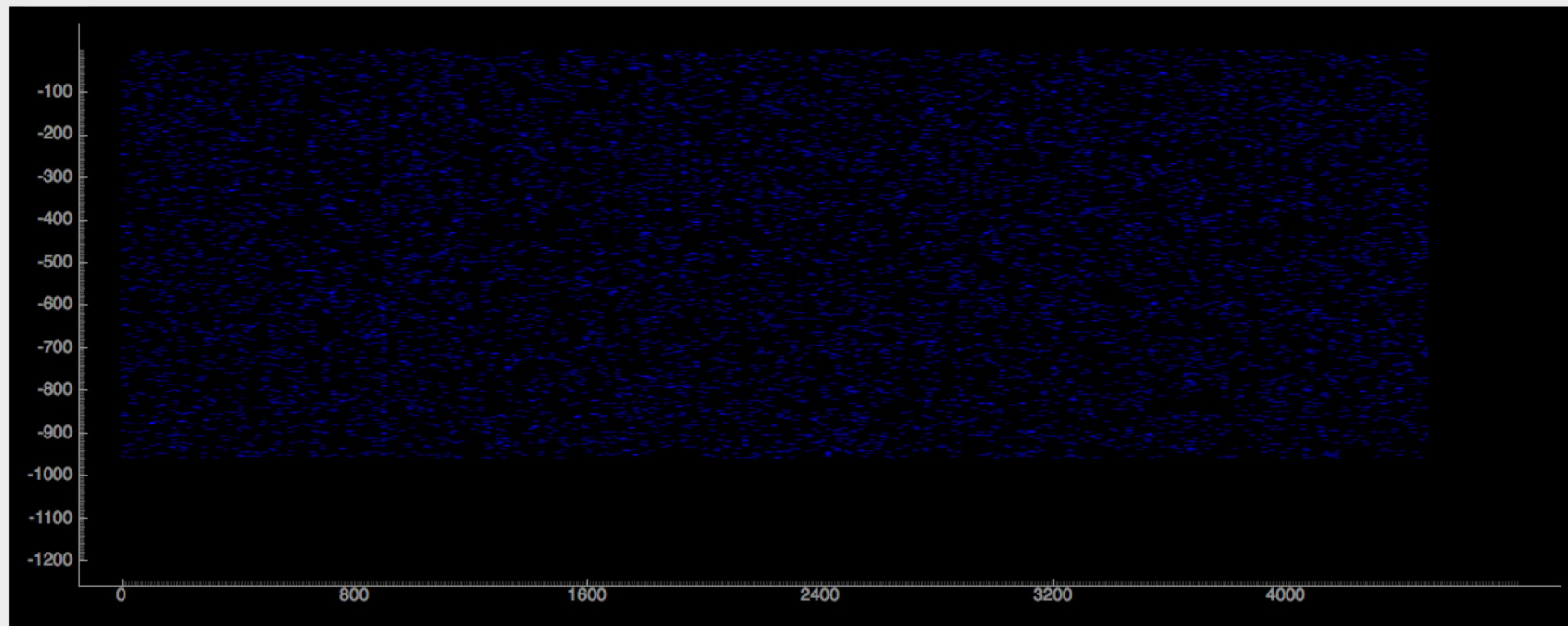


High-E(atmospheric)

E.g., background frame image input to CNN  
(before downsampling)

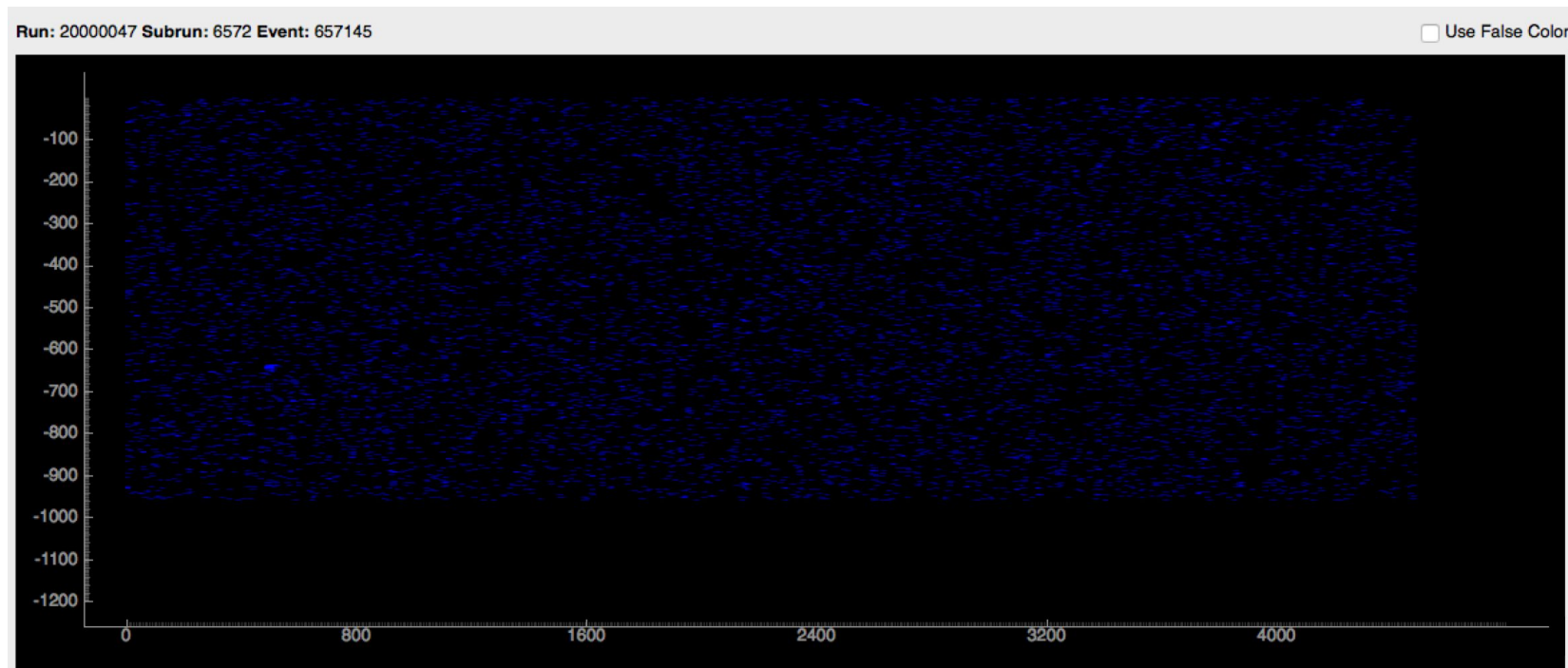
Run: 20000047 Subrun: 602 Event: 60106

☐ Use False Color



Radiologicals + noise

E.g., supernova frame image input to CNN  
(before downsampling)



SN

# Frame selection efficiency, requiring low RAD score

- From 3-class training of CNN (SN, background, and High-E)
- Classification definition: Keep events that satisfy RAD score cut.
- Efficiencies are shown separately on each exclusive frame type (SN, n-nbar, atmo. nu, p-decay, cosmic); note: only one interaction per frame assumed.

<b>RAD score cut</b>	<b>RAD frame efficiency</b>	<b>Data rate (RAD)</b>	<b>SN frame efficiency</b>	<b>n-nbar frame efficiency</b>	<b>atmo. nu frame efficiency</b>	<b>p-decay frame efficiency</b>	<b>cosmic frame efficiency</b>
<b>&lt;0.05</b>	0.56% (99.44% rejection)	6.4 GB/s (201 PB/year)	89%	100%	92%	99%	92%
<b>&lt;0.01</b>	0.18% (99.82% rejection)	2.05 GB/s (65 PB/year)	86%	100%	91%	99%	92%
<b>&lt;0.001</b>	0.031% (99.969% rejection)	350 MB/s (11 PB/year)	77%	100%	89%	98%	90%
<b>&lt;0.0002</b>	0.011% (99.989% rejection)	125 MB/s (3.9 PB/year)	69%	100%	87%	97%	88%

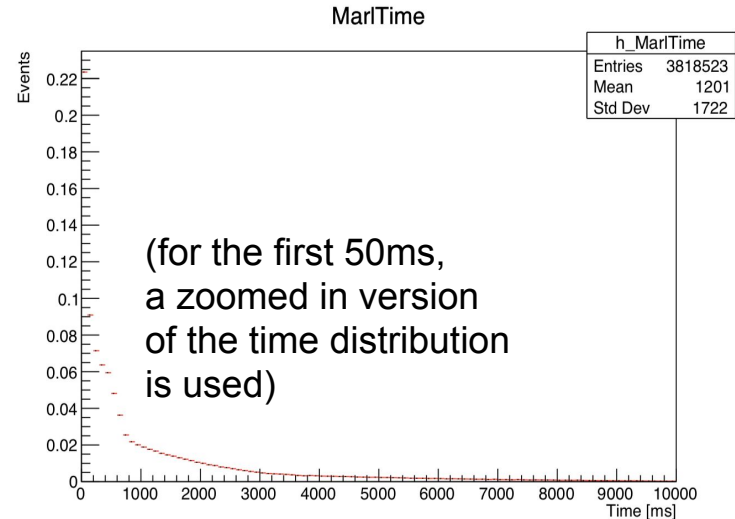
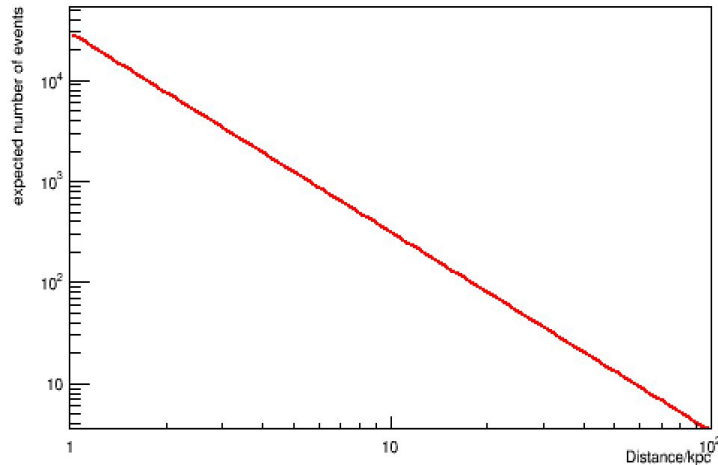
# CNN-based data selection in more detail

- Each **collection plane frame** is classified by a **CNN (VGG16b) network** trained on a set of frames containing **(1) supernova neutrino interactions**, **(2) high-energy off-beam interactions** (including atmospheric, n-nbar, cosmic, p decay), and **(3) radiological-only background**.
  - Each collection plane frame corresponds to 1APA x 1 drift time (960 collection plane wires x 2.25 ms x 2 MHz digitization, raw digits). Frames are downsampled to meet CNN input requirements (input image is 600x600 pixels) for the training and the inference.
  - Signal frames (SN, High-E) are defined as the ones containing true interaction vertex, but may only partially contain interaction final states.
- The network is trained to produce **3 scores for each frame**: RAD, SN, High-E. Frames are kept **according to their RAD score**. (We keep frames with very low RAD score.)

# CNN-based SN burst trigger study

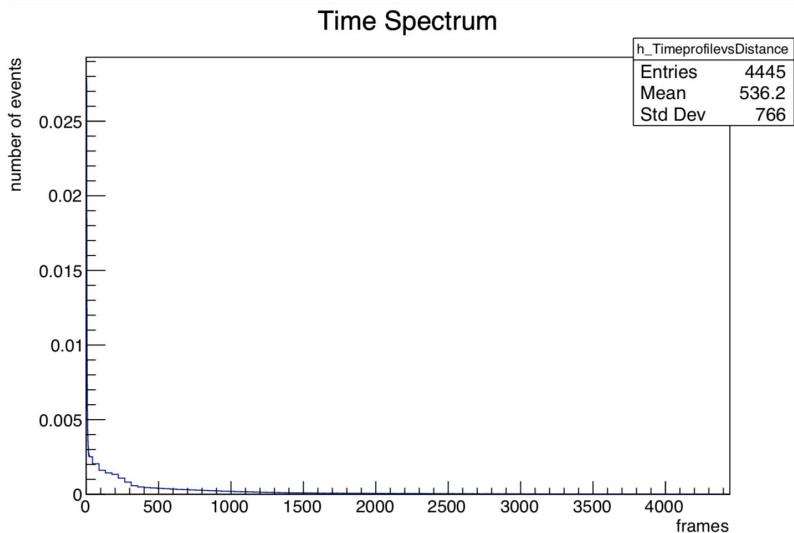
- Supernova timing profile lasts as long as 10s, which corresponds to 4445 frames. SN burst trigger aims to scan 200 collection plane drift volumes for 10 sec (4445 frames) to obtain better efficiency than ~69% of one frame per APA efficiency.
- SN neutrino event rate, timing profile (provided by Kate Scholberg; same as for traditional hit finding study; see the previous talk) is applied to a given distance, in order to obtain the event rate per frame for 10s (=4445 frames).

Expected number of events for a given supernova distance

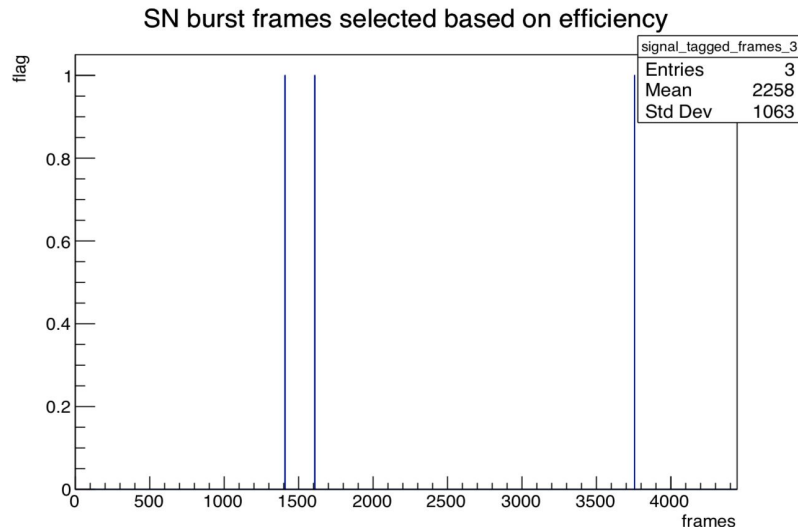


# CNN-based SN burst trigger study

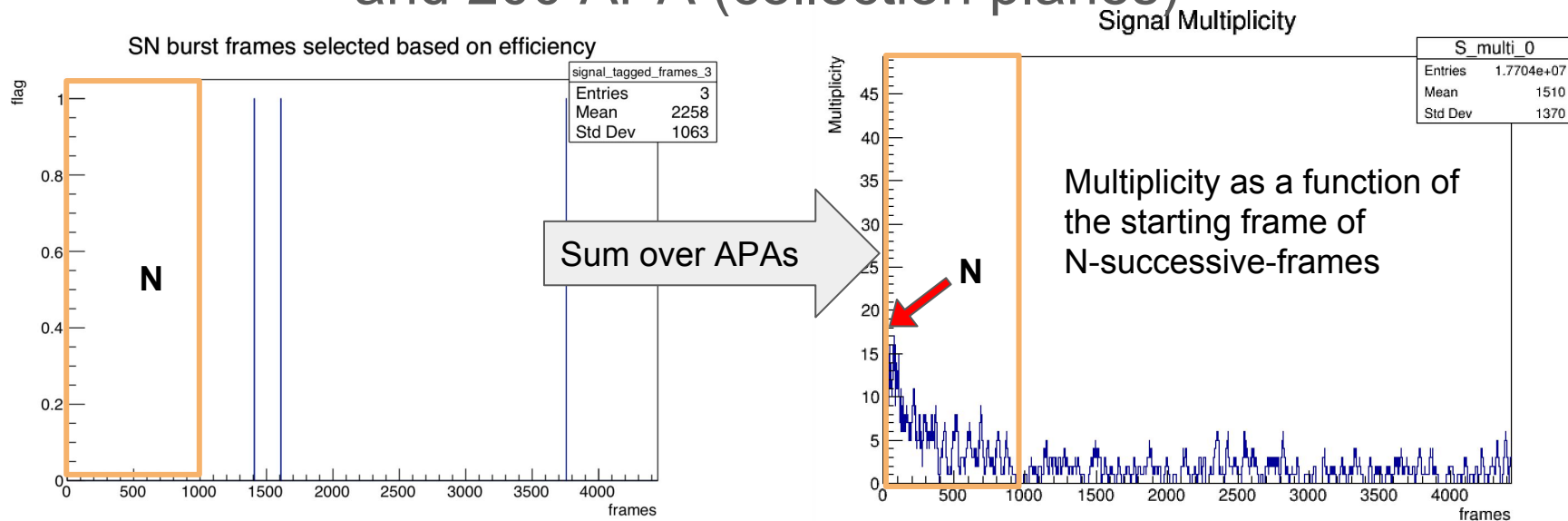
When each frame across 10s is filled with event distribution, Poisson fluctuations are applied and our selection efficiencies from CNN (68.84% for SN filled frame, 0.011% for empty frame), for both SN bursts and background.



1 frame=2.25ms



# Tagged frame multiplicity for a single APA (collection planes) and 200 APA (collection planes)

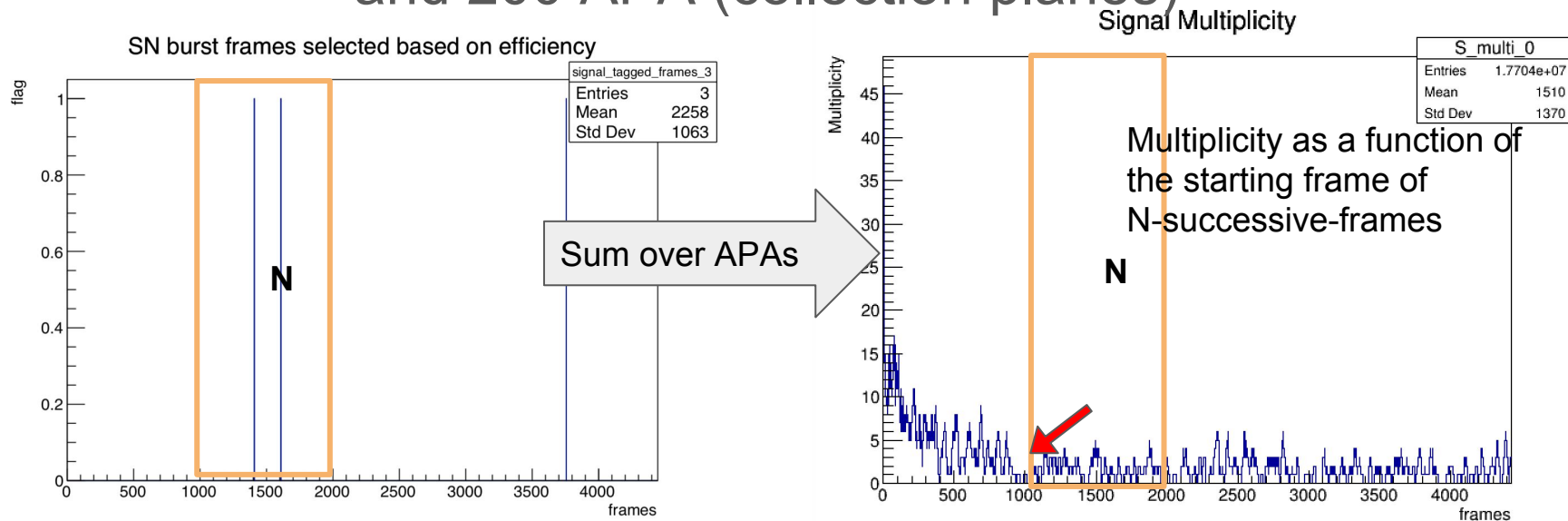


For a SN burst at 10kpc (plus background), a block size of N frames strides from 0s to 10s.

Over 200 APA collection planes, the multiplicity is calculated to be:  
tagged number of frames within the window of N-successive-frames over 200 APA collection plane frames.

N-successive-frames window strides from the first frame to the end.

# Tagged frame multiplicity for a single APA (collection planes) and 200 APA (collection planes)



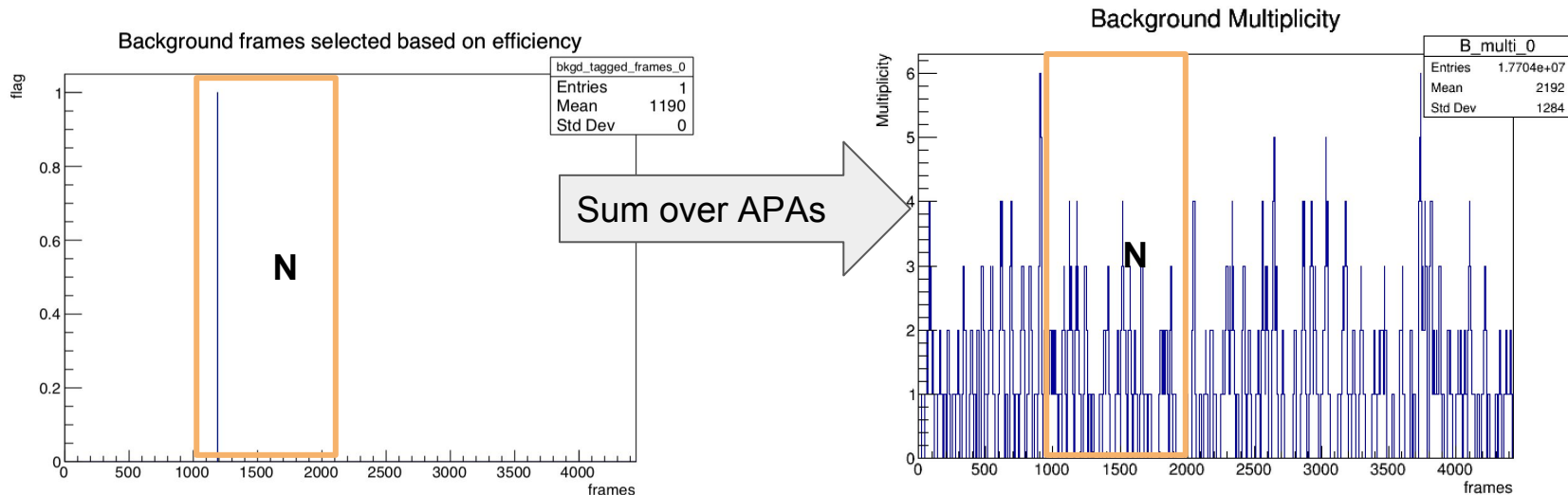
For a SN burst at 10kpc (plus background), a block size of N frames strides from 0s to 10s.

Over 200 APA collection planes, the multiplicity is calculated to be:

tagged number of frames within the window of N-successive-frames over 200 APA collection plane frames.

N-successive-frames window strides from the first frame to the end.

# Tagged frame multiplicity for a single APA (collection planes) and 200 APA (collection planes)

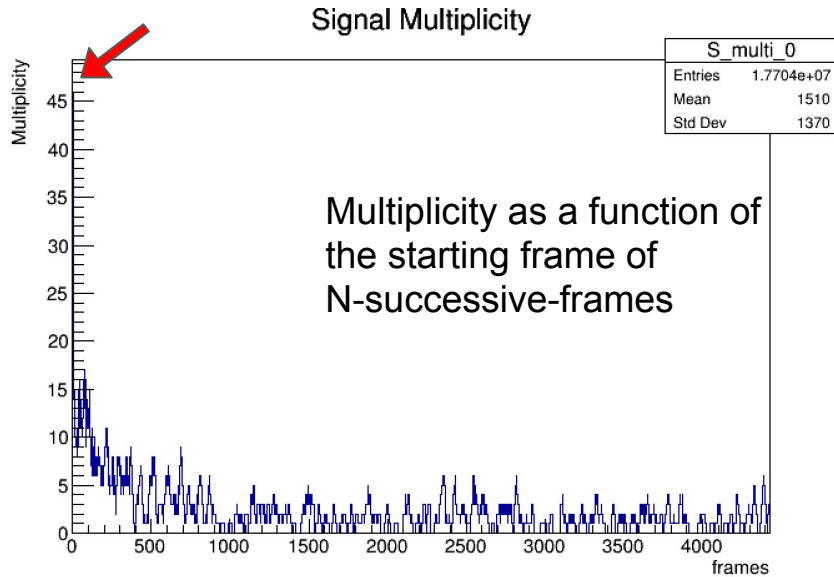


For a background, a block size of N frames strides from 0s to 10s.

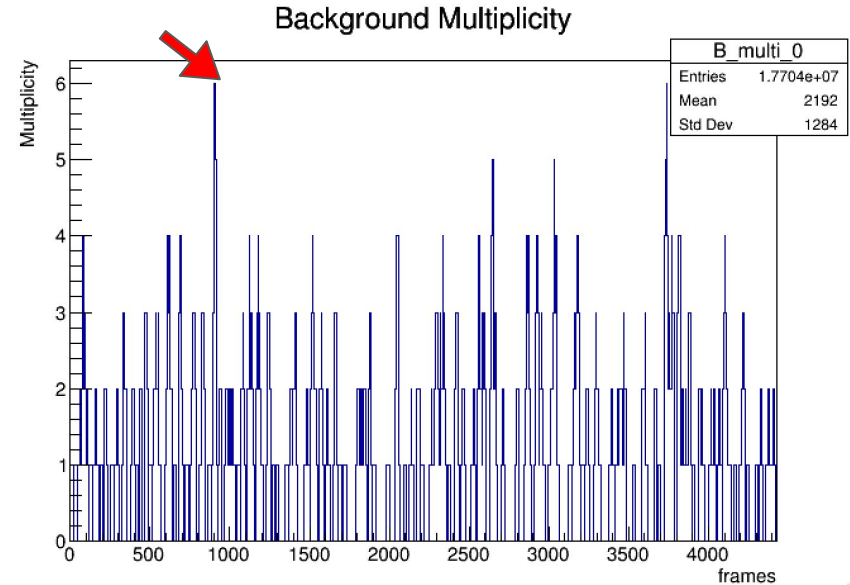
Over 200 APA collection planes, the multiplicity is calculated to be:  
tagged number of frames within the window of N successive frames over 200 APA collection plane frames.

N successive frame window strides from the first frame and to the end.

# Maximum multiplicities for signal and background for 10s X 200 APA collection planes



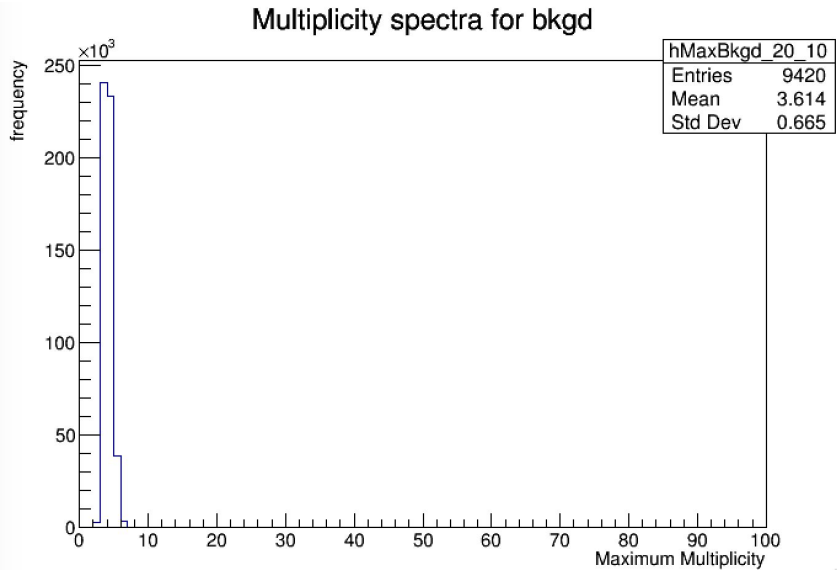
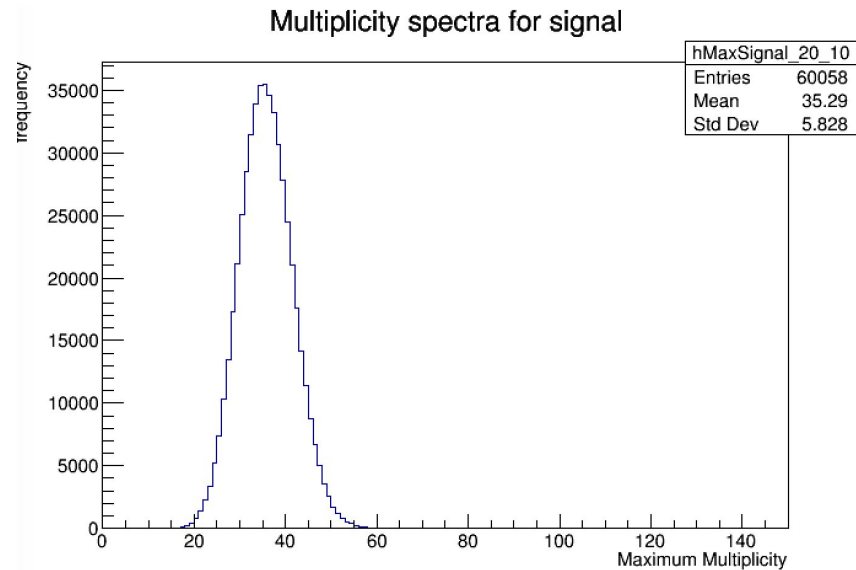
SN multiplicity over N=20 at 10kpc  
Maximum multiplicity of the SN burst : 45

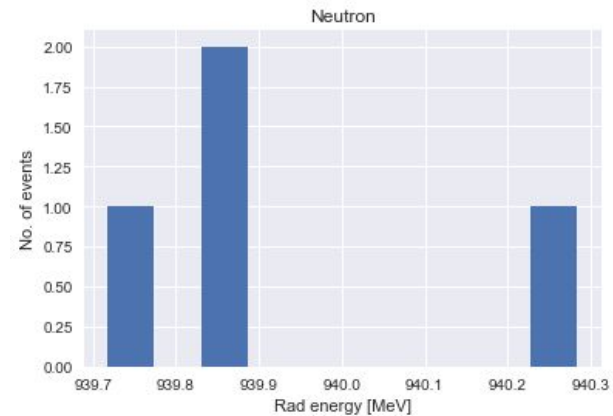
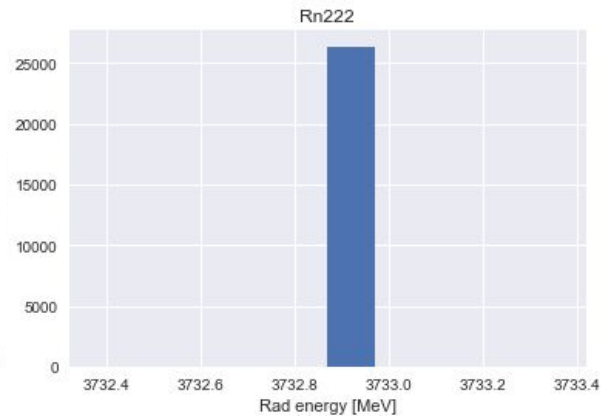
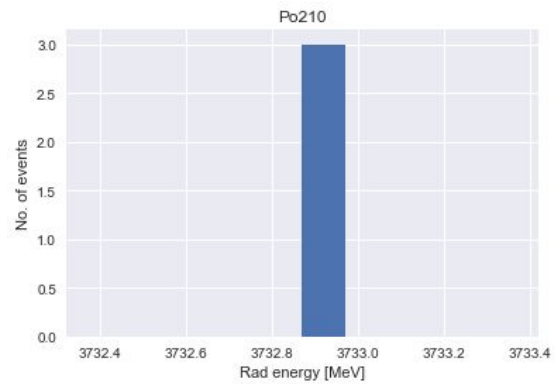


background multiplicity over N=20 at 10kpc  
Maximum multiplicity of the fake burst : 6

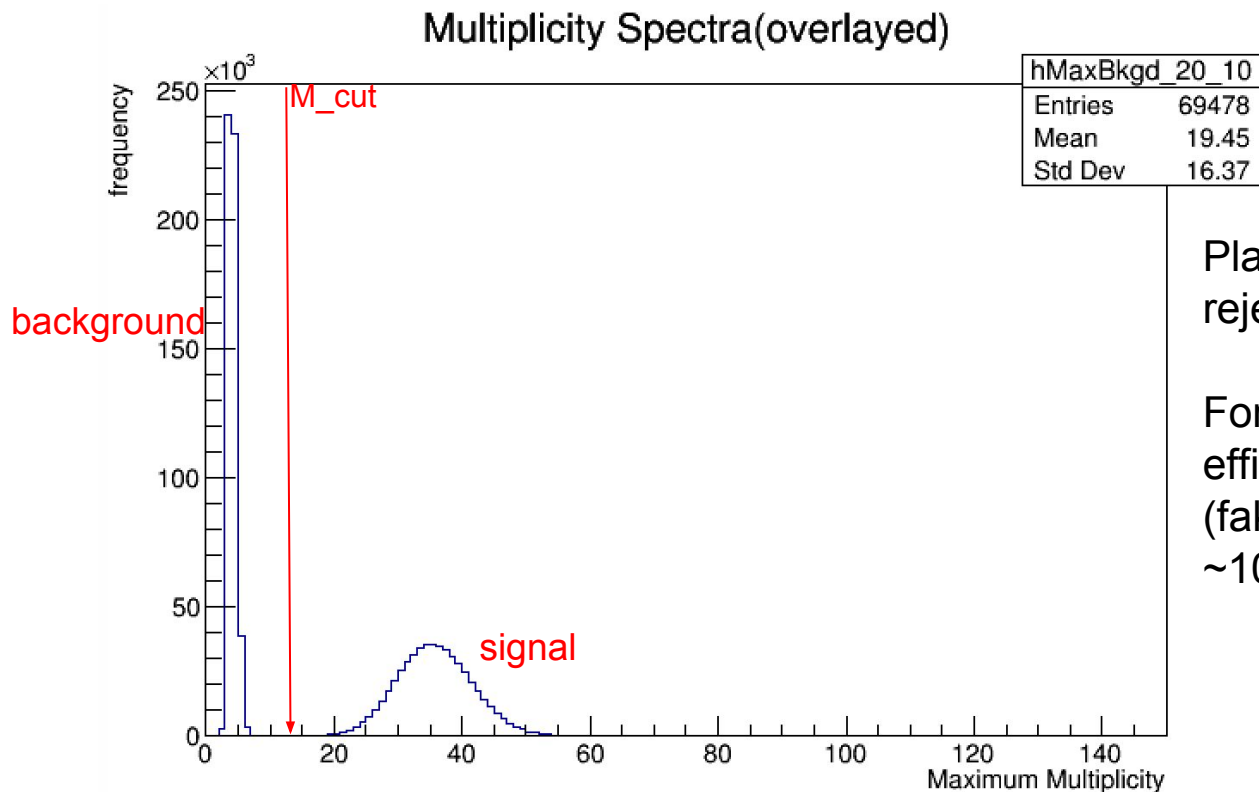
# Maximum Multiplicity Spectra: SN at L=10kpc

## Integration N=20 across 200 collection planes for 520k bursts





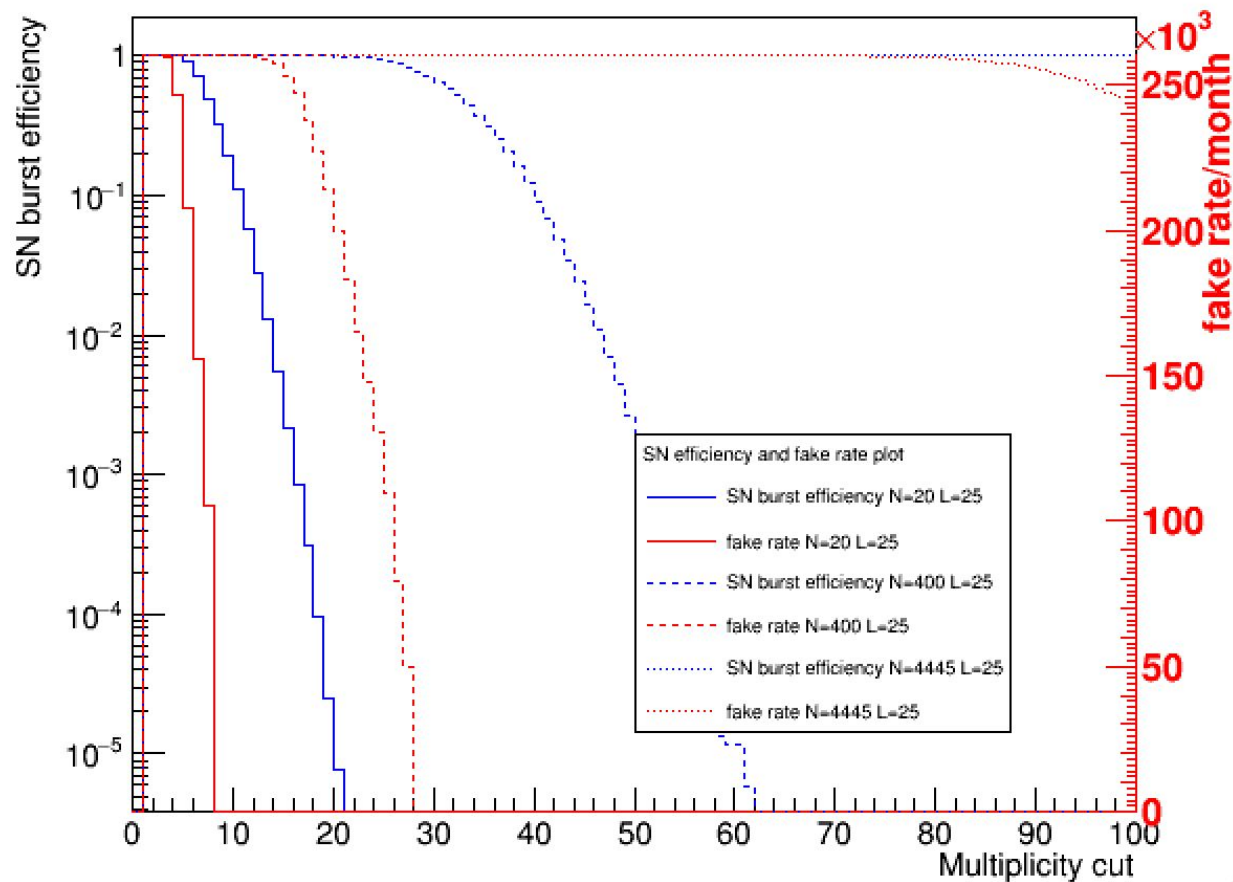
# Overlaid multiplicity spectra @ L=10kpc, N=20



Place  $M_{\text{cut}}$  where we can reject background.

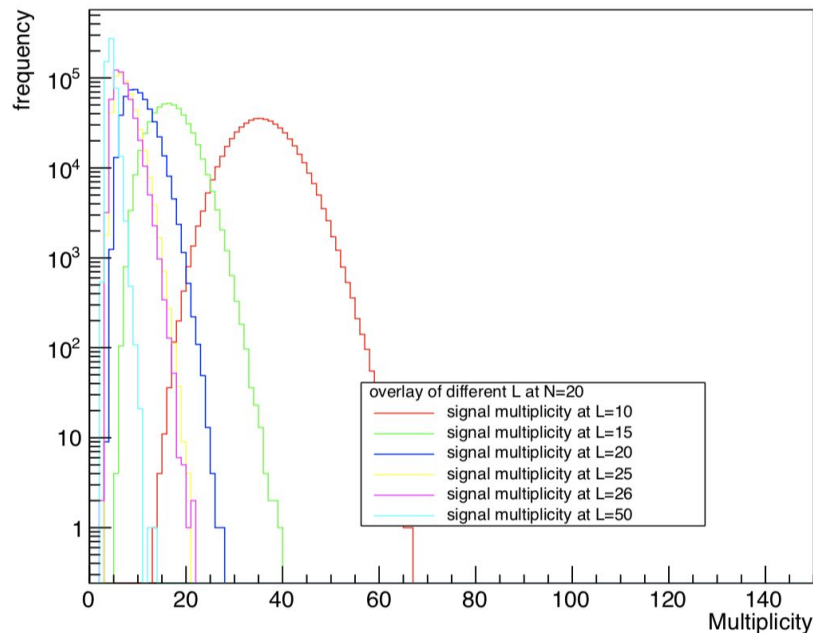
For  $L=10\text{kpc}$ ,  $N=20$ , we could efficiently cut background (fake rate  $\leq 1/\text{month}$ ) and get  $\sim 100\%$  SN burst efficiency.

SN burst efficiency & fake rate vs multiplicity cut L=25



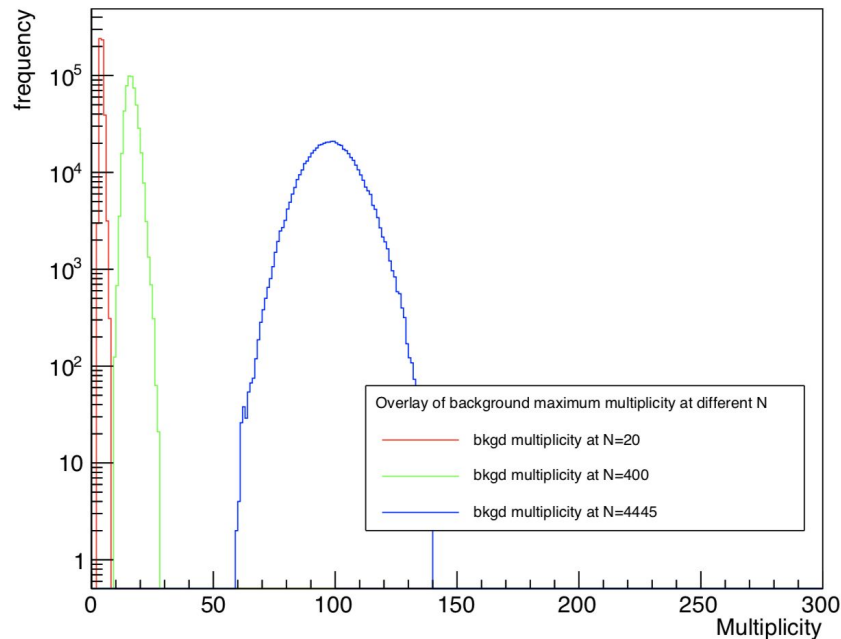
# Overlaid maximum multiplicity spectra

overlay of signal multiplicity at different L (N=20)



As distance gets larger, maximum multiplicity value gets smaller.

Overlay of background at different N



As N gets larger, maximum multiplicity value gets larger.